

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 27, 2003, 11:12:13 ; Search time 289 Seconds  
(without alignments)  
1.234 Million cell updates/sec

Title: al512306  
Perfect score: 133984  
Sequence: 1 GCAGGCC

RECEIVED 10/10/68

Gapop 10.0 , Gapext 0.5

Searched:	1 seqs, 1331 residues	2
Total number of hits satisfying chosen parameters:		

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					
Result No.	Score	Query Match	Length	ID	Description
C 1	720.4	0.5	1331	1	US-09-830224B-2
2	37.4	0.0	1331	1	US-09-830224B-2
					Sequence 2, Appl 1
					Sequence 2, Appl 1

## ALIGNMENTS

```

RESULT 1
US-09-830224B-2/c
; Sequence 2, Application US/09830244B
; GENERAL INFORMATION:
; APPLICANT: TANG, Y. Tom
; APPLICANT: CORLEY, Neil C.
; APPLICANT: GUEGLER, Karl J.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: LYSINE-RICH STATHERIN PROTEIN
; FILE REFERENCE: PF-0620 USN
; CURRENT APPLICATION NUMBER: US/09/830,244B
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: PCT/US99/24046
; PRIOR FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/155,209
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 1331
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 2820214CB1
; FEATURE:

```

```

QY 91841 GGAGATGGGTGAACCGGAGCGGAGGTTGAGTGAGCCGAGATCGCGCCACTGCAC 91900
Db 708 ----- 709
QY 91901 TCCAGCCGGGCGACAGAGAGAGACTCCGTCCTAAAAAAGAAATGGAGAGTT 91960
Db 708 -----TTAAAAAACAAGAAATGGAGAGTT 679
QY 91961 CTACTAACTGCTGCTCCAGAGAGGAGATCAATGAGGATCGTAAGAGAAACAAAC 92020
Db 678 CTACTAACTGCTGCTCCAGAGAGGAGATCAATGAGGATCGTAAGAGAAACAAAC 619
QY 92021 CTGAGGGGAGGAGGAGATCGTGAATTCAGATAAGTACCAAAATATATGCTATCT 92080
Db 618 CTGAGGGGAGGAGGAGATCGTGAATTCAGATAAGTACCAAAATATATGCTATCT 559
QY 92081 TTCTGAGAGATCCCTCTCTCTGTTAAACATTTCAAGTCAATCAAAAGGCTATGTTAG 92140
Db 558 TTCTGAGAGATCCCTCTCTCTGTTAAACATTTCAAGTCAATCAAAAGGCTATGTTAG 499
QY 92141 AGGTGAGCAAAATCCCAAGTGTAGATTTACTGTAACAGTTTATTAAGTATGCTGATTC 92200
Db 498 AGGTGAGCAAAATCCCAAGTGTAGATTTACTGTAACAGTTTATTAAGTATGCTGATTC 439
QY 92201 CCAATACGGCATTTACCAATATGTCATTTAGTGCACATCAAAAGATCTGAAATGCTAAT 92260
Db 438 CCAATACGGCATTTACCAATATGTCATTTAGTGCACATCAAAAGATCTGAAATGCTAAT 379
QY 92261 GGGACATTTCCCAAAAGGAAAGAAATCCATGCTCTATTAAGGGAGATATAGCAA 92320
Db 378 GGGACATTTCCCAAAAGGAAAGAAATCCATGCTCTATTAAGGGAGATATAGCAA 319
QY 92321 TGGCAAGCAGATGAATTTCTGTCAAAAGAAAGTACTCTCAAGCAATCAGGAAATA 92380
Db 318 TGGCAAGCAGATGAATTTCTGTCAAAAGAAAGTACTCTCAAGCAATCAGGAAATA 259
QY 92381 CTGTTCCCAATTTAGGCTGGGAGCCACATAGCCATCTCAGATATGATTTATTT 92440
Db 258 CTGTTCCCAATTTAGGCTGGGAGCCACATAGCCATCTCAGATATGATTTATTT 199
QY 92441 GTTAGCCCTTAGGGCCCTCAATATTTCCCCCTCATCTGTAACCTCACTGTTACTCA 92500
Db 198 GTTAGCCCTTAGGGCCCTCAATATTTCCCCCTCATCTGTAACCTCACTGTTACTCA 139
QY 92501 CAAGTCTGATTAATTCAGCAAGTAACTTAACATTCCTGATGCGAGTGCAGAAATCCA 92560
Db 138 CAAGTCTGATTAATTCAGCAAGTAACTTAACATTCCTGATGCGAGTGCAGAAATCCA 79
QY 92561 GACATTTGTAACACATGTAATTTTCAAGTTTGGTACATGAGACAGCTTACCGGTTATCT 92620
Db 78 GACATTTGTAACACATGTAATTTTCAAGTTTGGTACATGAGACAGCTTACCGGTTATCT 19
QY 92621 TTGATTTCTAAACATAG 92638
Db 18 TTGATTTCTAAACATAG 1

```

RESULT 2

```

US-09-830224B-2
; Sequence 2, Application US/09830244B
; GENERAL INFORMATION:
; APPLICANT: TANG, Y. Tom
; APPLICANT: CORLEY, Neil C.
; APPLICANT: GUEGLER, Karl J.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: LYSINE-RICH STATHERIN PROTEIN
; FILE REFERENCE: PF-0620 USN
; CURRENT APPLICATION NUMBER: US/09/830,244B
; PRIOR FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: PCT/US99/24046
; PRIOR FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/155,209
; PRIOR FILING DATE: 1998-10-23

```

```

; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 1331
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: 2820214CB1
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 1242
; OTHER INFORMATION: a, t, c, g, or other
US-09-830224B-2

```

```

Query Match 0.08; Score 37.4; DB 1; Length 1331;
Best Local Similarity 87.24; Pred. No. 0;
Matches 41; Conservative 0; Mismatches 6; Indels 0; Caps 0;
QY 70292 AATGCTTGGGAGACTAATTTGAGTAATAATAAACTCCGGTCTCCC 70338
Db 912 AACTTTTGGGAGACTGATTTGAGTAATAATAAACTCTGGTCTCCC 958

```

Search completed: June 27, 2003, 11:17:21  
Job time : 301 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model  
Run on: June 27, 2003, 12:14:57 ; Search time 13.5 Seconds  
(without alignments)  
1.886 Million cell updates/sec

Title: al512306  
Perfect score: 238409  
Sequence: 1 GCAGGCCAAGTCTCTGGCA.....AAACAAATGCTGAATTC 133984

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1 seqs, 95 residues

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL-frame\_n2p.model -DEV-soft -Q-al512306.seq -DB-mayes.pep -SUFFIX-ptp  
-OUT-mayesaa -MINMATCH-0.1 -LOOPEXT-0 -LOOPEXT-0 -UNITS-bits -START-1 -END--1  
-MATRIX-blosum62 -TRANS-human40.cdi -LIST-45 -DOALIGN-200 -THR\_SCORE-pct  
-THR\_MAX-100 -THR\_MIN-0 -ALIGN-15 -MODS-LOCAL -OUTFMT-ptp -NORM-ext  
-HEAPSIZE-500 -MINLEN-0 -MAXLEN-2000000000 -NCFU-6 -NO\_XLPXY -NEG\_SCORES-0  
-LONGLOG -THREADS-1 -XGAPOP-10 -XGAPEXT-0.5 -FGAPOP-6 -FGAPEXT-7 -YGAPOP-10  
-YGAPEXT-0.5 -DELOP-6 -DELEXT-7

Database : mayes.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query	Score	Match	Length	DB ID	Description
c 1	349	0.1	95	1	US-09-830244B-1	Sequence 1, Appli
2	58.5	0.0	95	1	US-09-830244B-1	Sequence 1, Appli

## ALIGNMENTS

RESULT 1  
US-09-830244B-1  
; Sequence 1, Application US/09830244B  
; GENERAL INFORMATION:  
; APPLICANT: TANG, Y. Tom  
; APPLICANT: CORLEY, Neil C.  
; APPLICANT: GUEGLER, Karl J.  
; APPLICANT: PATTERSON, Chandra  
; TITLE OF INVENTION: LYSINE-RICH STATHERIN PROTEIN  
; FILE REFERENCE: PF-0620 USN  
; CURRENT APPLICATION NUMBER: US/09/830,244B  
; PRIOR FILING DATE: 2002-12-13  
; PRIOR APPLICATION NUMBER: PCT/US99/24046  
; PRIOR FILING DATE: 1999-10-22  
; PRIOR APPLICATION NUMBER: 60/155,209

; PRIOR FILING DATE: 1998-10-23  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PERL Program  
; SEQ ID NO 1

; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: Incyte ID No: 2820214CD1  
US-09-830244B-1

## Alignment Scores:

Pred. No.: 0  
Score: 349.00  
Percent Similarity: 100.00%  
Best Local Similarity: 98.48%  
Query Match: 0.15%  
DB: 1  
Length: 95  
Matches: 65  
Conservative: 1  
Mismatch: 0  
Indels: 0  
Gaps: 0

al512306 (1-133984) x US-09-830244B-1 (1-95)

Qy 91100 ATGTGTTTCATAAAGTGGGAAGAAACACATTTAAAGTAACCTTTTGGGAGACTGAT 91041  
Db 1 MetTrpPheHisLysValGlyArgLysGlnHisPheLysValThrPheTrpGluThrAsp 20  
Qy 91040 TTGAGTAATAATAAACTCTGGTCTCCCTTAAGAAAAAAACCCCTTCCACCTTTACTGT 90981  
Db 21 LeuSerAsnAsnLysThrLeuValSerLeuLysLysLysProPheHisLeuTyrCys 40  
Qy 90980 GTCATTATATATCCCTTAGTTCACAAAGTTAATATCTTATTTCTGGATATTCCTTTTATA 90921  
Db 41 ValileTyrileProLeuValProLysLeullelleLeuPheLeuAspIleAlaPheille 60  
Qy 90920 CCAAGACCCCTTATCAGC 90903  
Db 61 ProlYsserLeulleSer 66

## RESULT 2

US-09-830244B-1  
; Sequence 1, Application US/09830244B  
; GENERAL INFORMATION:  
; APPLICANT: TANG, Y. Tom  
; APPLICANT: CORLEY, Neil C.  
; APPLICANT: GUEGLER, Karl J.  
; APPLICANT: PATTERSON, Chandra  
; TITLE OF INVENTION: LYSINE-RICH STATHERIN PROTEIN  
; FILE REFERENCE: PF-0620 USN  
; CURRENT APPLICATION NUMBER: US/09/830,244B  
; PRIOR FILING DATE: 2002-12-13  
; PRIOR APPLICATION NUMBER: PCT/US99/24046  
; PRIOR FILING DATE: 1999-10-22  
; PRIOR APPLICATION NUMBER: 60/155,209  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PERL Program  
; SEQ ID NO 1  
; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: Incyte ID No: 2820214CD1  
US-09-830244B-1

## Alignment Scores:

Pred. No.: 0  
Score: 58.50  
Percent Similarity: 47.54%  
Best Local Similarity: 32.79%  
Query Match: 0.02%  
DB: 1  
Length: 95  
Matches: 20  
Conservative: 9  
Mismatch: 24  
Indels: 8  
Gaps: 2

```

a1512306 (1-133984) x US-09-830244B-1 (1-95)
QY 104105 AAGAAATGTTACCCCTTCAAAAGAAAAAACCACTTCCCTTCAAAATGTGAATTTAC--- 104161
Db 25 LysThrLeuValSerLeuLysLysLysLysProPheHisLeuTyrCysValIleTyrIle 44
QY 104162 -----CTGAATTTGTCTGTAATCAAGTTGAGCCATTACTTCTAGGTGTATA 104206
Db 45 ProLeuValProLysLeuIleLeuPheLeuAspIleAlaPheIleProLysSerLeu 64
QY 104207 GTTCTTCTCAAGTTTCTTAAACCAAGCAGCCAGCCAGCTACATCCCACTCTCTCAA 104266
Db 65 Ile---SerGln-PheGlnAsnAsnHisTyrThrHisAsnHisThrAsnHisAsnThrAs 83
QY 104267 T 104267
Db 83 n 83

```

Search completed: June 27, 2003, 12:16:25  
Job time : 39.5 secs